

# Anticipate & Act: Supplementary Material

## Contents

<b>1 Activities</b>	<b>2</b>
1.1 Regular Activities . . . . .	2
1.1.1 Illustrative Example . . . . .	2
1.2 Special Sequence: Weekly Routine . . . . .	3
1.2.1 Illustrative Example . . . . .	3

# 1 Activities

## 1.1 Regular Activities

We created a dataset  $\mathcal{T}$  of high-level tasks in a household environment. These tasks belong to activities such as *Cooking*, *Cleaning*, *Washing*, *Baking*, and *Gardening*. We then generated a set of task routines  $\mathcal{R}_i$ , each with  $\approx 20$  tasks, by sampling tasks across different activities while preserving the relative order of tasks within each activity. We follow a pipelining approach while creating the routines, where long duration tasks are pipelined with shorter duration tasks. This methodology allowed us to create task routines spanning activities of daily living in a home. The dataset of activities is defined as follows:

- **Cooking:** Get a vegetable, Clean the vegetable, Cut the vegetable, Cook the vegetable, Serve food, Take out leftover food
- **Washing clothes:** Collect dirty clothes, Load dirty clothes in the washing machine, Run the washing machine, Collect clean clothes, Dry the clean clothes, Iron the dry clothes, Fold the clothes, Put away the clean clothes
- **Cleaning the house:** Collect cleaning supplies, Dust the surfaces, Vacuum the floors, Put away the cleaning supplies, Take out the trash
- **Washing dishes:** Collect dirty dishes, Load dirty dishes in the dishwasher, Run the dishwasher, Put away the clean dishes
- **Gardening:** Cut the grass, Water the plants, Put away the gardening tools
- **Baking:** Preheat the oven, Prepare baking ingredients, Pour the batter into the baking pan, Bake in the oven, Serve cake

Here, we measure the accuracy of any LLM to complete tasks belonging to a given activity, while maintaining the relative order of tasks within an activity.

### 1.1.1 Illustrative Example

Let's examine a specific instance:

Input: Collect dirty dishes, Get a vegetable

The Large Language Model (LLM) generates the following sequence of tasks:

LLM Output: Collect dirty dishes, Load dirty dishes in the dishwasher, Run the dishwasher, Get a vegetable, Clean the vegetable, Put away the clean dishes, Cut the vegetable, Cook the vegetable, Serve food, Take out leftover food.

This sequence showcases the LLM's ability to not only follow the order of input tasks but also logically complete each activity. Notably, it recognizes that 'Run the dishwasher' is a long-duration task and intelligently pipelines tasks from other activities (in this case, tasks related to 'Cooking') during this period. This allows for efficient task management, and we see that LLMs are able to understand this pattern of pipelining.

## 1.2 Special Sequence: Weekly Routine

We also created a dataset  $W$  of weekly routines. These routines are a sequence of tasks that are performed on specific days of the week. The dataset is defined as follows: Sure, here's how you can complete the list of weekly tasks in LaTeX:

- **Monday:** Cook omelette, Serve the food, Collect dirty dishes, Load dirty dishes in the dishwasher, Run the dishwasher, Put away the clean dishes, Dust the surfaces, Cut the grass, Water the plants, Put away the gardening tools
- **Tuesday:** Cook cereal, Serve the cereal, Collect dirty clothes, Load dirty clothes in the washing machine, Run the washing machine, Collect clean clothes, Dry the clean clothes, Iron the dry clothes, Fold the clothes, Put away the clean clothes, Collect dirty dishes, Load dirty dishes in the dishwasher, Run the dishwasher, Put away the clean dishes
- **Wednesday:** Cut a mango, Cut an apple, Serve the fruits, Prepare a coffee, Serve the coffee, Collect cleaning supplies, Dust the surfaces, Vacuum the floors, Put away the cleaning supplies, Take out the trash, Collect dirty clothes, Load dirty clothes in the washing machine, Run the washing machine
- **Thursday:** Get a vegetable, Clean the vegetable, Cut the vegetable, Cook the vegetable, Serve food, Prepare milk, Serve milk, Take out leftover food, Collect dirty dishes, Load dirty dishes in the dishwasher, Run the dishwasher, Put away the clean dishes, Cut the grass, Water the plants, Put away the gardening tools

There are some extra tasks are as follows: Extinguish the fire, Set up laptop, Repair the washing machine, Repair the dishwasher, Repair the oven, Serve wine, Clean windows

In this case, we measure how accurately an LLM can complete tasks for a given day while considering any specific prompt given by the user. The prompts by user change the usual routine for a day by either removing some tasks, or replacing some existing tasks with a new task.

### 1.2.1 Illustrative Example

Let's examine a specific instance:

Input: Today is Monday, I have an urgent meeting in the morning.

The LLM outputs as following:

LLM Output: Cook omelette, Serve the food, Set up Laptop, Collect dirty dishes, Load dirty dishes in the dishwasher, Run the dishwasher, Put away the clean dishes, Dust the surfaces, Cut the grass, Water the plants, Put away the gardening tools

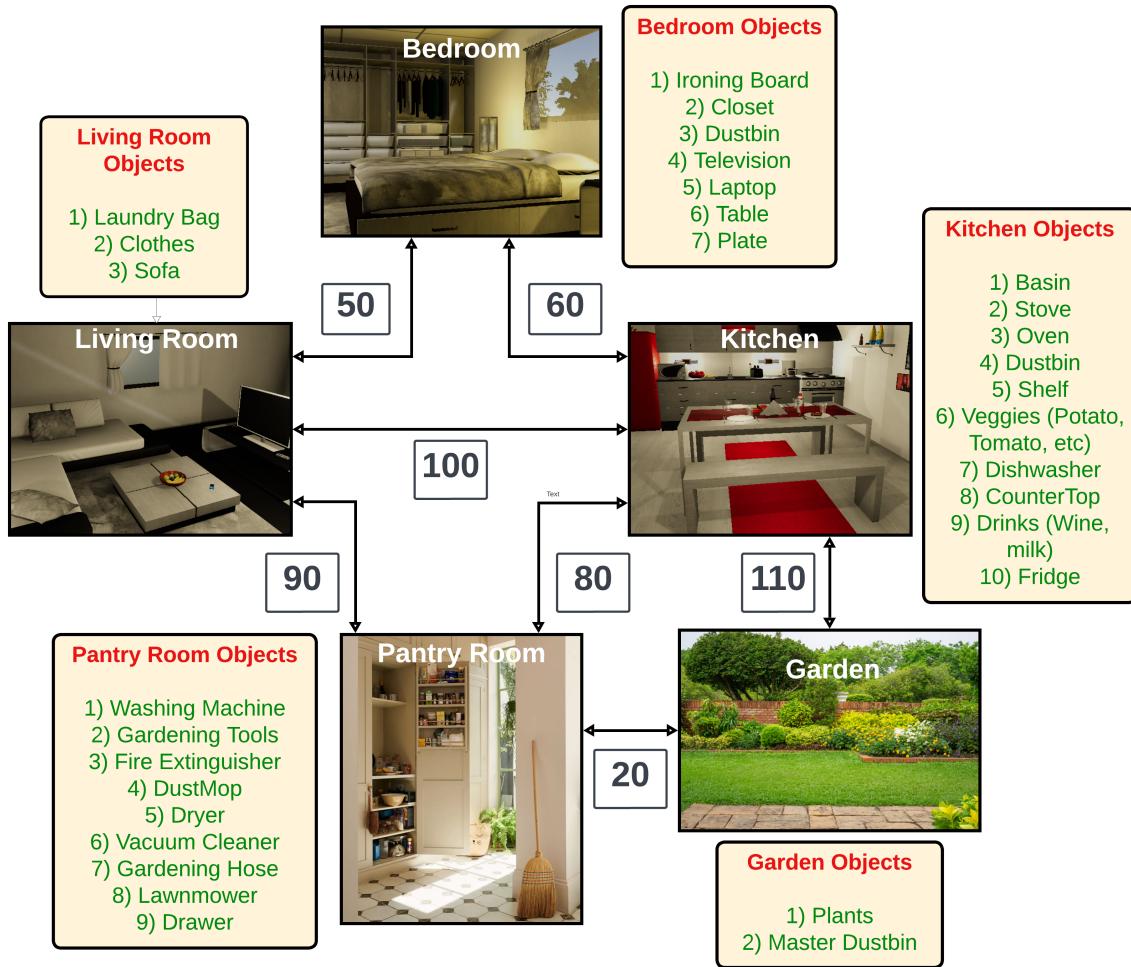


Figure 1: Room Map

## **References**